



A markov random field model-based approach to unsupervised texture segmentation using local and global spatial statistics

Charles Kervrann, Fabrice Heitz

► To cite this version:

Charles Kervrann, Fabrice Heitz. A markov random field model-based approach to unsupervised texture segmentation using local and global spatial statistics. [Research Report] RR-2062, INRIA. 1993. inria-00074610

HAL Id: inria-00074610

<https://inria.hal.science/inria-00074610>

Submitted on 24 May 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Unité de recherche INRIA Lorraine, Technôpole de Nancy-Braboïs, Campus scientifique,
615 rue de Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unité de recherche INRIA Rennes, IRISA, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unité de recherche INRIA Rhône-Alpes, 46 avenue Félix Viallet, 38031 GRENOBLE Cedex 1
Unité de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105,
78153 LE CHESNAY Cedex
Unité de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS
Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
(France)
ISSN 0249-6399

- [42] C.S. WON and H. DERIN. – Unsupervised segmentation of noisy and textured images using Markov Random Fields. – *CVGIP : Graphical Models and Image Processing*, Vol. 54, No 4: pages 308–328, July 1992.
- [43] Z. WU and R. LEAHY. – A new unsupervised hierarchical segmentation algorithm for textured images. – In *Proc. Int. Conf. Acoust., Speech, Signal Processing*, pages 2325–2328, Albuquerque, 1990.
- [44] A.L. YUILLE. – Feature extraction from faces using deformable templates. – *Int. J. Computer Vision*, Vol. 8, No 2: pp. 99–111, 1992.
- [45] J. ZHANG and J.W. MODESTINO. – Unsupervised image segmentation using a Gaussian model. – In *Proc. CISS 1988*, pages 7–10, Princeton, NJ, 1988.

- [32] B. JULESZ. – Experiments in the visual perception of textures. – *Scientific American*, No 232: pages 34–43, April 1975.
- [33] A. KHOTANZAD and J.Y. CHEN. – Unsupervised segmentation of textured images by edge detection in multidimensional features. – *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 1, No 4: pages 414–421, 1989.
- [34] J. KONRAD and E. DUBOIS. – Bayesian estimation of motion vector fields. – *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 14, No 9: pages 910–927, 1992.
- [35] S. LAKSHMANAN and H. DERIN. – Simultaneous parameter estimation and segmentation of Gibbs random fields using simulated annealing. – *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 11, No 8: pages 799–813, 1989.
- [36] B.S. MANJUNATH and R. CHELLAPPA. – Unsupervised texture segmentation using Markov Random Fields. – *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 13: pages 478–482, 1991.
- [37] J.W. MODESTINO and J. ZHANG. – A Markov Random Field model-based approach to image interpretation. – *IEEE Trans. Pattern Anal. Machine Intell.*, Vol 14, No 6: pages 606–615, June 1992.
- [38] H.H. NGUYEN and P. COHEN. – Gibbs random fields, fuzzy clustering, and the unsupervised segmentation of textured images. – *CVGIP : Graphical Models and Image Processing*, Vol. 55, No 1: pages 1–19, January 1993.
- [39] F. PRETEUX and X. DESCOMBES. – Texture synthesis and analysis : multiscale process cooperation (in french). – In *8e congrès AFCET/INRIA Reconnaissance des Formes et Intelligence Artificielle*, pages 1015–1022, Lyon, France, 1991.
- [40] J. RISSANEN. – Modeling by shortest data description. – *Automatica*, Vol 14: pages 465–472, 1986.
- [41] J.F. SILVERMAN and D.B. COOPER. – Bayesian clustering for unsupervised estimation of surface and texture models. – *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 10, No 4: pages 482–495, July 1988.

-
- Pattern Anal. Machine Intell.*, Vol. 6, No 6: pages 721–741, November 1984.
- [22] S. GEMAN, D. GEMAN, and C. GRAFFIGNE. – Locating texture and object boundaries. – In P.A. Devijver and J. Kittler, editors, *Pattern Recognition. Theory and Applications*, volume NATO ASI Series, Vol. F30, pages 165–177, Heidelberg, 1987. Springer Verlag.
 - [23] C. GRAFFIGNE. – *Experiments in Texture Analysis and Segmentation*. – PhD thesis, Brown University, 1986.
 - [24] R.M. HARALICK. – Statistical and structural approaches to texture. – *Proc. IEEE*, Vol. 67, No 5: pages 786–804, May 1979.
 - [25] R.M. HARALICK. – Statistical image texture analysis. – In T.S Young and K.S. Fu, editors, *Handbook of Pattern Recognition*, pages 247–279. Academic Press, 1986.
 - [26] R.M. HARALICK, K. SHANMUGAM, and ITS’HAK DINSTEIN. – Textural features for image classification. – *IEEE Trans. Systems, Man, Cyber.*, Vol. 3, No 6: pages 610–621, Nov. 1973.
 - [27] F. HEITZ and P. BOUTHEMY. – Multimodal estimation of discontinuous optical flow using Markov Random Fields. – *IEEE Trans. Pattern Anal. Machine Intell.*, accepted for publication (also available as Technical Report INRIA No 1367, 1991), 1993.
 - [28] F. HEITZ and C. KERVRANN. – A statistical model-based approach for unsupervised texture segmentation. – In *8th Scandinavian Conf. Image Analysis*, pages 281–288, Tromso, Norway, May 1993.
 - [29] F. HEITZ, P. PEREZ, and P. BOUTHEMY. – Multiscale minimization of global energy functions in some visual recovery problems. – *CVGIP : Image Understanding*, accepted for publication, 1993.
 - [30] J.Y. HSIAO and A.A. SAWCHUK. – Unsupervised textured image segmentation using feature smoothing and probabilistic relaxation techniques. – *Comput. Vision, Graphics, Image Processing*, Vol. 48: pages 1–21, 1989.
 - [31] R. HU and M.M. FAHMY. – Texture segmentation based on a hierarchical Markov random field model. – *Signal Processing*, Vol. 26: pages 285–305, 1992.

-
- [11] P.B. CHOU and C.M. BROWN. – The theory and practice of bayesian image modeling. – *Int. J. Comp. Vis.*, Vol 4: pages 185–210, 1990.
 - [12] F.S. COHEN and Z. FAN. – Maximum likelihood unsupervised texture segmentation. – *CVGIP : Graphical Models and Image Processing*, Vol. 54, No 3: pages 239–251, 1992.
 - [13] R. CONNERS, M. TRIVEDI, and C. HARLOW. – Segmentation of high-resolution urban scene using texture operators. – *Comput. Vision, Graphics, Image Processing*, Vol. 25: pages 273–310, 1984.
 - [14] G.R. CROSS and A.K. JAIN. – Markov Random Field texture models. – *IEEE Trans. Pattern Anal. Machine Intell.*, Vol 5, No 1: pages 25–39, January 1983.
 - [15] H. DERIN and H. ELLIOT. – Modeling and segmentation of noisy and textured images using Gibbs random fields. – *IEEE Trans. Pattern Anal. Machine Intell.*, Vol 9, No 1: pages 39–55, Jan. 1987.
 - [16] M. DERRAS, M. BERDUCAT, and P. BONTON. – Vision-guided mower for the upkeep of the natural environment. – In *Proc. Int. Seminar : Use of On-Machine Vision Systems for the Agricultural and Bio-industries*, pages 37–46, Montpellier, France, Sept. 1991.
 - [17] R.C. DUBES and A.K. JAIN. – Random field models in image analysis. – *J. Applied Statistics*, Vol. 16, No 2: pages 131–163, 1989.
 - [18] P.W. FUNG, G. GREBBIN, and Y. ATTIKIOUZEL. – Model-based region growing segmentation of textured images. – In *Proc. Int. Conf. Acoust., Speech, Signal Processing*, pages 2313–2316, Albuquerque, USA, April 1990.
 - [19] M. GALLOWAY. – Texture analysis using gray level run lenghts. – *Comput. Vision, Graphics, Image Processing*, Vol 4: pages 171–199, 1974.
 - [20] D. GEMAN, S. GEMAN, C. GRAFFIGNE, and D. PONG. – Boundary detection by constrained optimization. – *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 12, No 7: pages 609–628, July 1990.
 - [21] S. GEMAN and D. GEMAN. – Stochastic relaxation, Gibbs distributions and the bayesian restoration of images. – *IEEE Trans.*

References

- [1] S.T. BARNARD. – Stochastic stereo matching over scale. – *Int. J. Comp. Vis.*, Vol 3: pages 17–32, 1989.
- [2] M. BASSEVILLE, A. BENVENISTE, K.C. CHOU, S.A. GOLDEN, R. NIKOUKHAH, and A.S. WILLSKY. – Modeling and estimation of multiresolution stochastic processes. – *IEEE Trans. Information Theory*, Vol. 38, No 2: pages 766–784, 1992.
- [3] J. BESAG. – On the statistical analysis of dirty pictures. – *J. Royal Statist. Soc.*, Vol. 48, Serie B, No 3: pages 259–302, 1986.
- [4] J.R. BEVERIDGE, J. GRIFFITH, R. KOHLER, A. HANSON, and E. RISEMAN. – Segmenting images using localized histograms and region merging. – *Int. J. Computer Vision*, Vol. 2: pages 311–347, 1989.
- [5] P. BOUTHEMY and E. FRANÇOIS. – Motion segmentation and qualitative dynamic scene analysis from an image sequence. – *Int. J. Computer Vision*, Vol 10, No 2: pages 157–182, 1993.
- [6] P. BOUTHEMY and P. LALANDE. – Recovery of moving object masks in an image sequence using local spatio-temporal contextual information. – *Optical Engineering*, Vol.32, No 6: pages 1205–1212, June 1993.
- [7] P. BRODATZ. – *Textures : A Photographic Album for Artists and Designers*. – Dover Publications, 1966.
- [8] S. CHATTERJEE. – Classification of natural textures using Gaussian Markov random field models. – In R. Chellappa and A. Jain, editors, *Markov Random Fields - Theory and application*, pages 159–177. Academic Press, 1993.
- [9] R. CHELLAPPA and S. CHATTERJEE. – Classification of textures using Gaussian Markov Random Fields. – *IEEE Trans. Acoust., Speech, Signal Processing*, Vol. 33, No 4: pages 959–963, Aug. 1985.
- [10] J.-L. CHEN and A. KUNDU. – Automatic unsupervised texture segmentation using Hidden Markov Model. – In *Proc. Int. Conf. Acoust., Speech, Signal Processing*, pages V.21–V.24, Minneapolis, USA, April 1993.

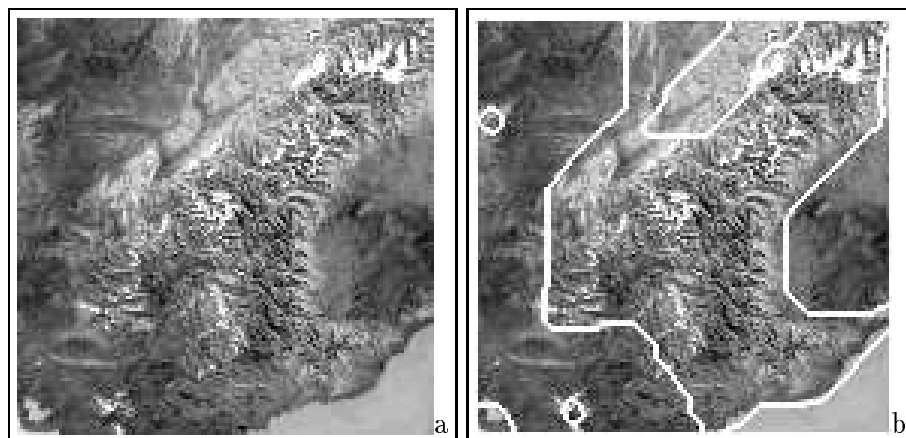


Figure 12: (a) satellite image of the Alps (France), (b) result of the unsupervised segmentation procedure (the boundaries of the different regions are superimposed on the original image) ($p = 0.25$)

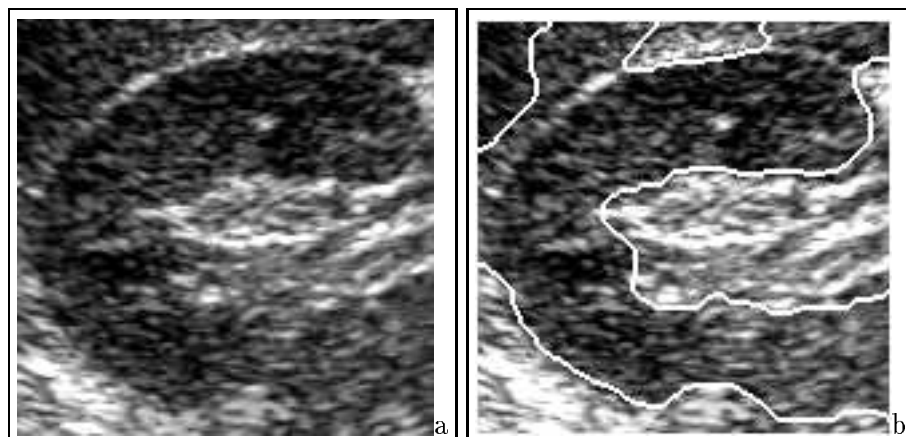


Figure 13: (a) Ultrasound image of a kidney (by courtesy of LEP-GDR134), (b) Result of the unsupervised segmentation procedure (the boundaries of the different regions are superimposed on the original image) ($p = 0.3$)

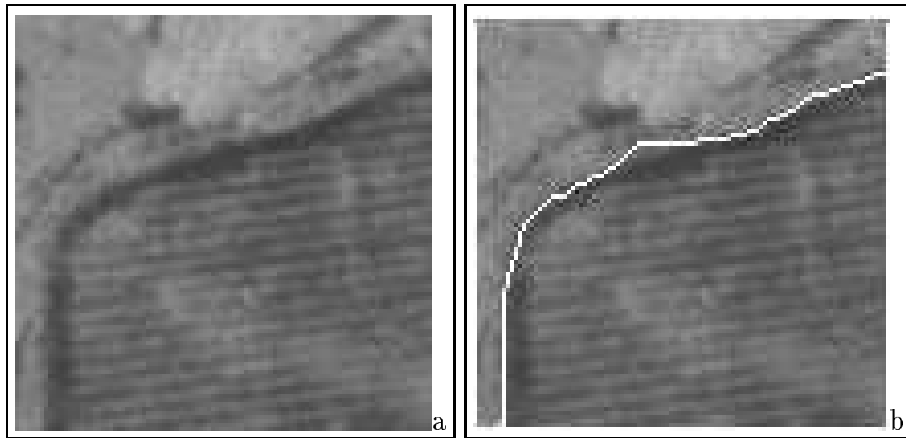


Figure 10: (a) Aerial image (Calvi area, by courtesy of DRET - Matra mission 1987) , (b) Result of the unsupervised segmentation procedure (the boundaries of the different regions are superimposed on the original image) ($p = 0.2$)

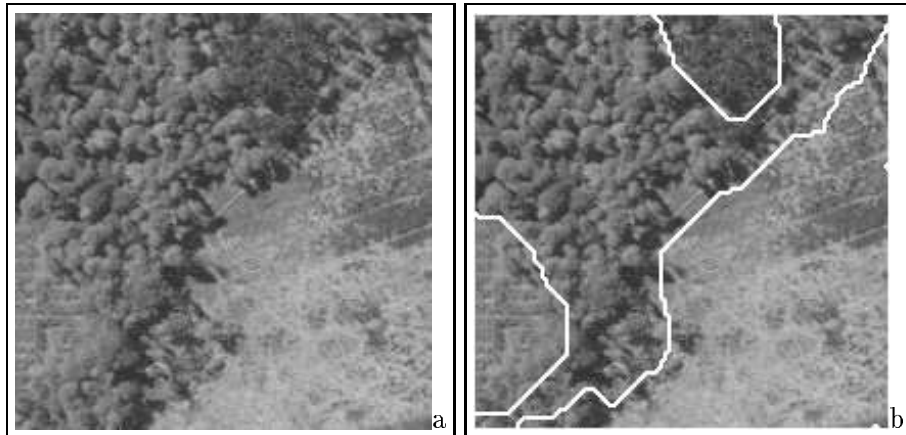


Figure 11: (a) Aerial image (Roumard area, by courtesy of DRET - INRA mission 1976), (b) Result of the unsupervised segmentation procedure (the boundaries of the different regions are superimposed on the original image) ($p = 0.65$)

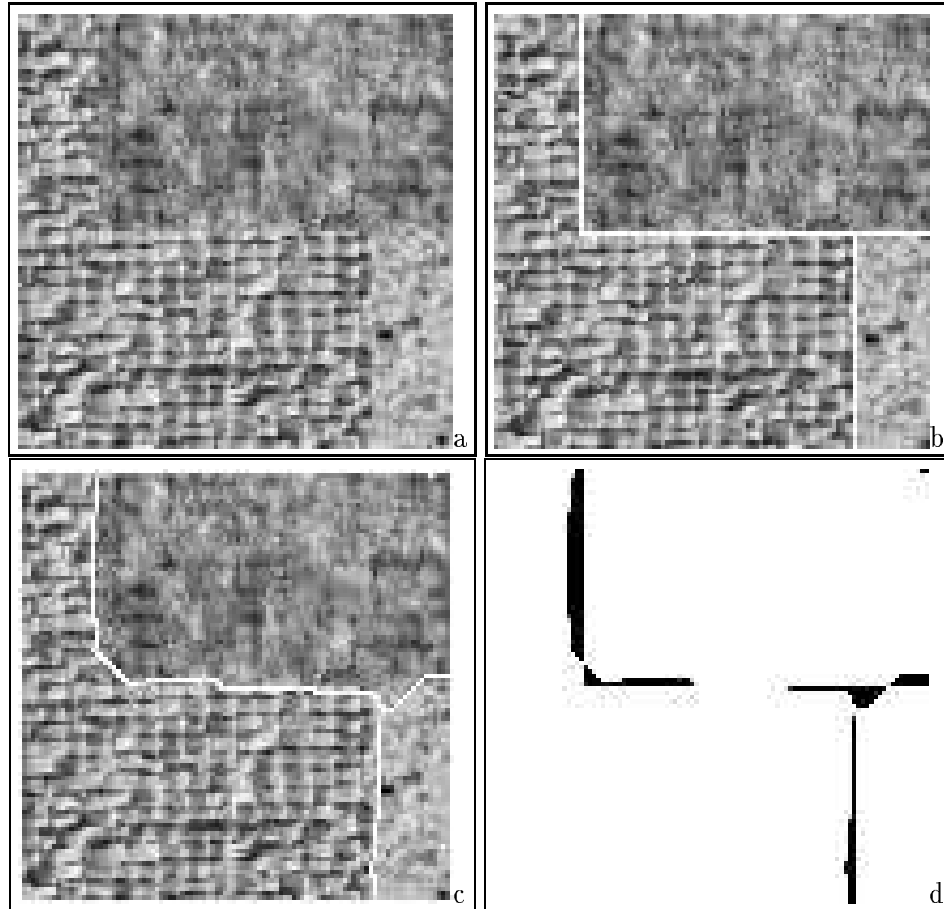


Figure 9: (a) 3-region textured image – (b) Ground truth (the boundaries of the different regions are superimposed on the original image) – (c) Result of the unsupervised segmentation procedure ($p = 0.65$) – (d) Difference between the ground truth and the unsupervised segmentation

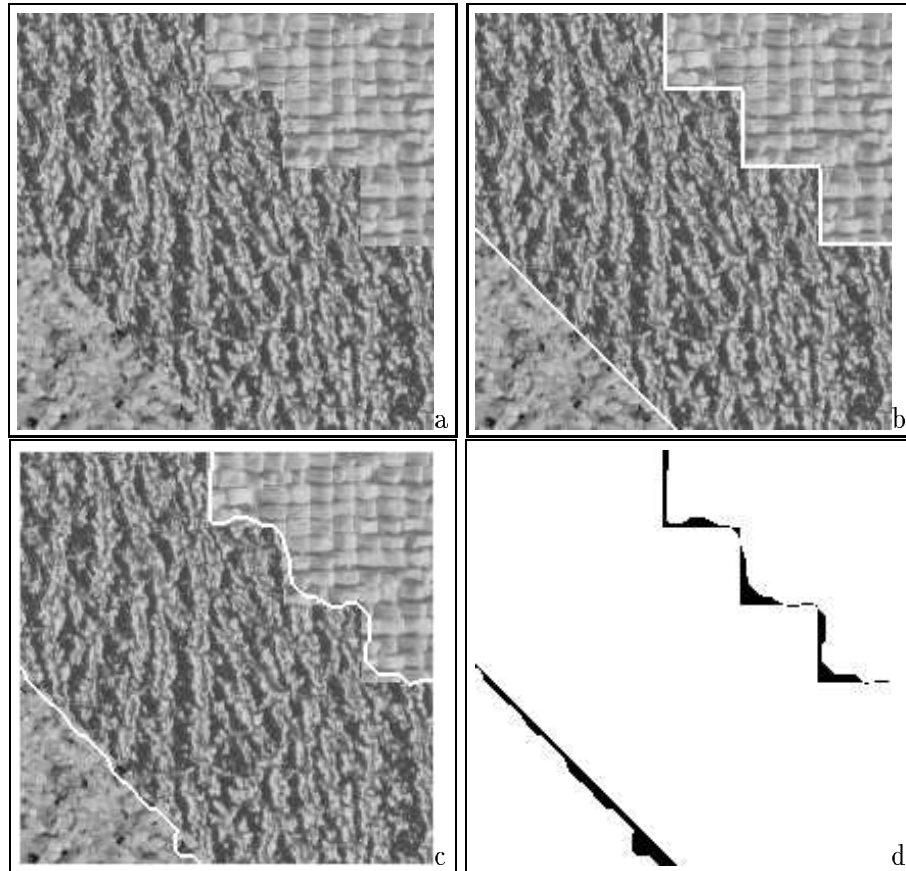


Figure 8: (a) 3-region textured image – (b) Ground truth (the boundaries of the different regions are superimposed on the original image) – (c) Result of the unsupervised segmentation procedure ($p = 0.75$) – (d) Difference between the ground truth and the unsupervised segmentation

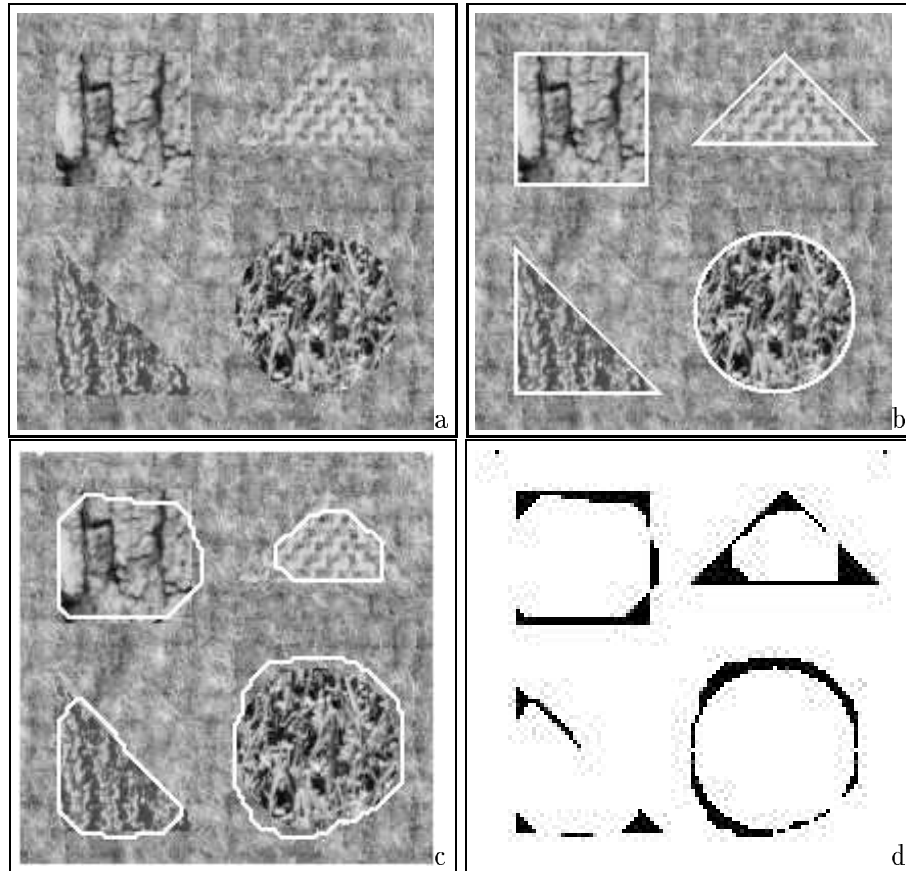


Figure 7: (a) 5-region textured image – (b) Ground truth (the boundaries of the different regions are superimposed on the original image) – (c) Result of the unsupervised segmentation procedure ($p = 0.65$) – (d) Difference between the ground truth and the unsupervised segmentation

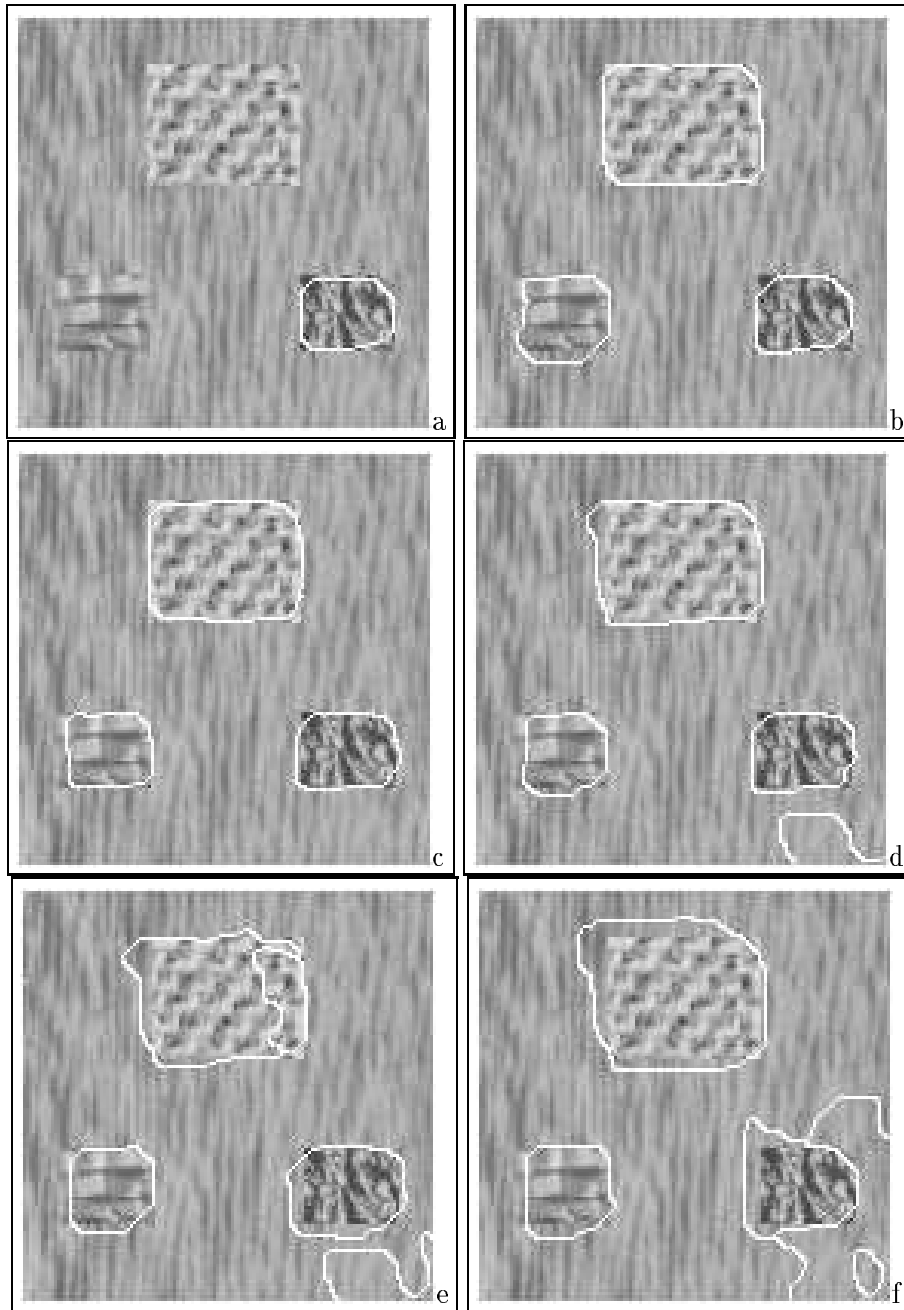


Figure 6: *Sensitivity of the final segmentation to the adjustment of parameter p : (a) $p = 0.3$, (b) $p = 0.4$, (c) $p = 0.5$, (d) $p = 0.6$, (e) $p = 0.7$, (f) $p = 0.8$.*

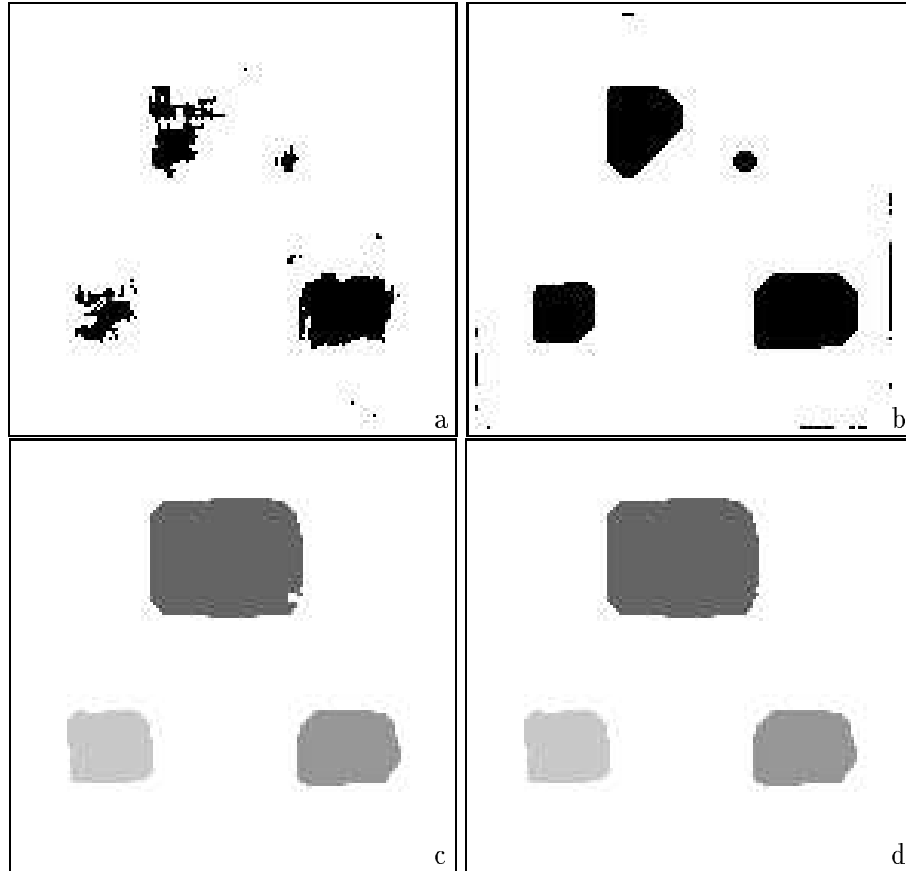


Figure 5: *Intermediate steps of the relaxation process : (a) initial maximum likelihood estimate, (b) first relaxation step (black points correspond to sites labeled ρ , (c) intermediate relaxation step (after relabeling of connected components) , (d) final segmentation*

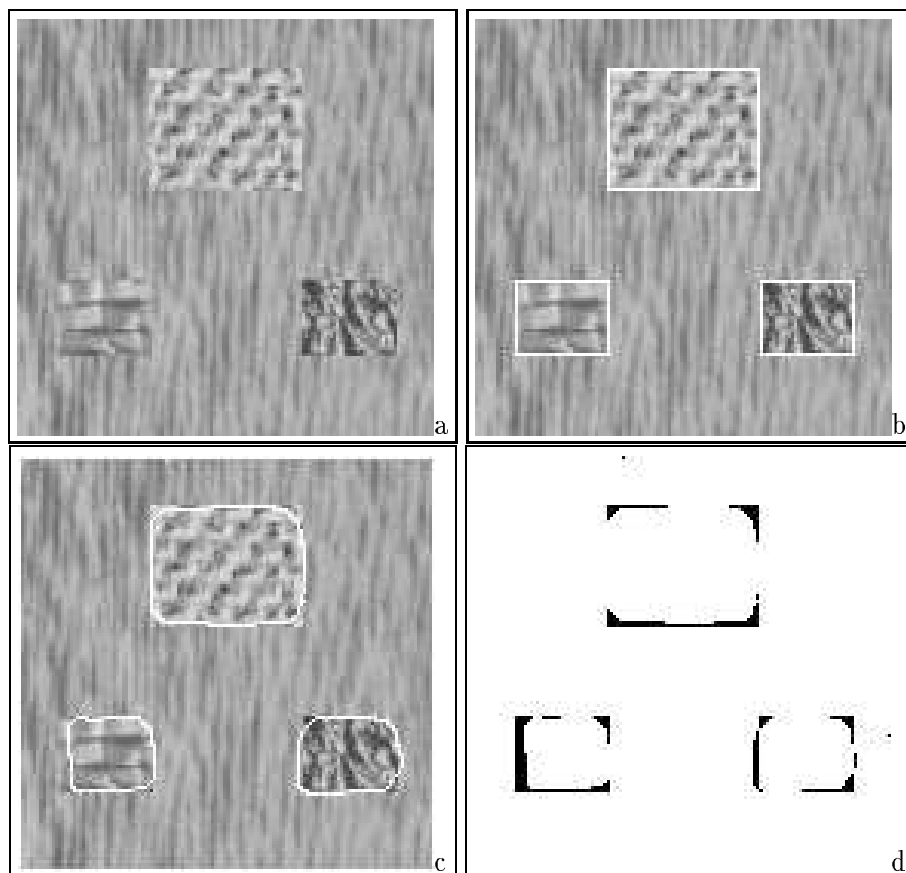
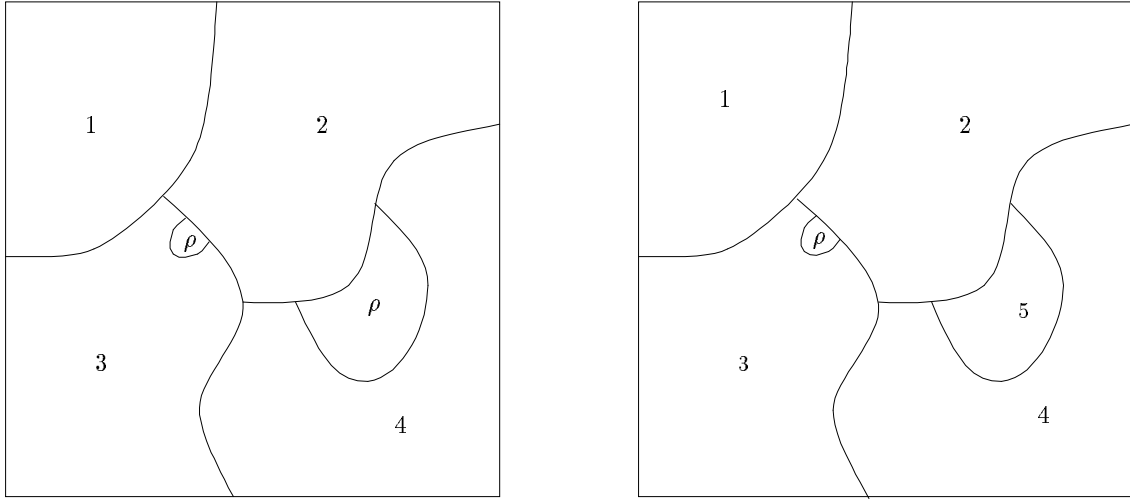


Figure 4: (a) 4-region textured image – (b) Ground truth (the boundaries of the different regions are superimposed on the original image) – (c) Result of the unsupervised segmentation procedure ($p = 0.5$) – (d) Difference between the ground truth and the unsupervised segmentation



before relabeling : $\hat{R} = 4$

after relabeling : $\hat{R} = 5$

Figure 3: *Relabeling of the outlier class ρ . Label ρ is maintained on small connected components (of size less than 1 % of the image size). New regions are created for all other connected components*

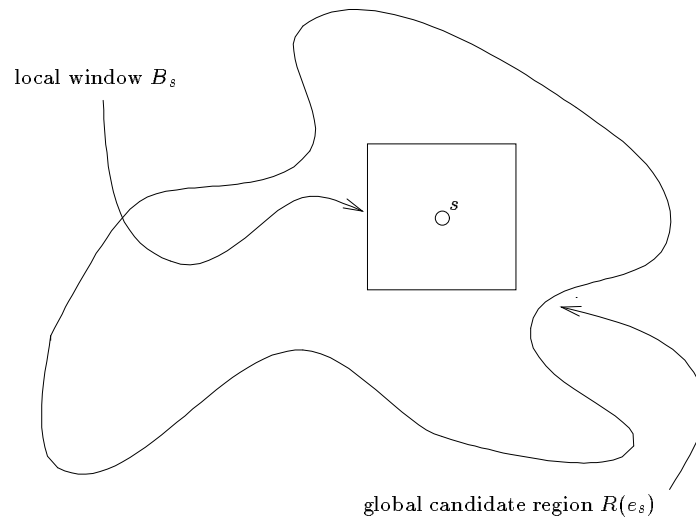


Figure 2: *Sets of texture features supporting the local and global statistical distributions*

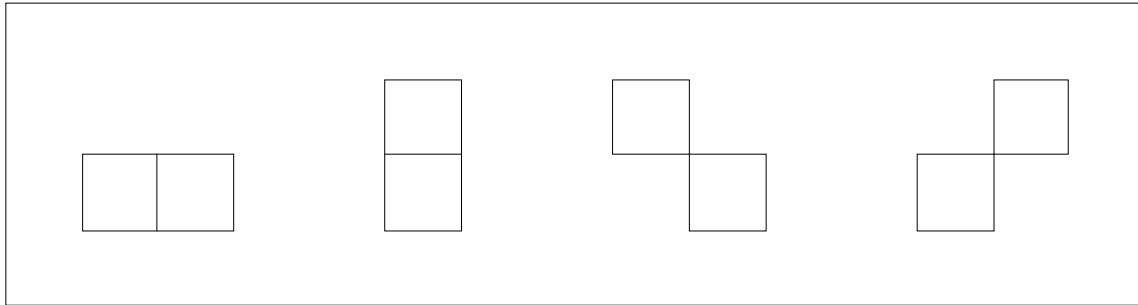


Figure 1: *Two element (doubleton) cliques associated to the a priory energy term U_2*

our knowledge, this problem remains an open issue, in an *unsupervised* framework.

Besides, the images processed in the examples of this paper were mainly composed of micro-textures. The processing of macro-textures would require significant alterations to the existing method. Multiscale estimation methods and random processes [2, 29, 39], for instance, appear particularly promising as far as the fusion of multiscale texture features in the final decision is concerned.

Acknowledgements

This work has been supported by MRT and CNRS in the context of the PRC Program "Man Machine Interface" and by the GDR TdSI 134.

ject of interest from the background. A deviation with respect to the actual boundary is observed in the upper part of the kidney, in which the texture pattern is very similar to the background. High-level model-guided segmentations, using “deformable templates” [44] for instance, would of course be better adapted to this class of images, for which a strong a priori knowledge about the object structure is usually available.

These different examples demonstrate the ability of the method to produce relevant segmentations on real images, even when the texture classes and the number of regions are not known a priori. Let us emphasize again that the only parameter that has been tuned in all our experiments was the probability p . Of course, an estimation procedure for p (from the data) would be necessary for a completely data-driven segmentation scheme.

Figures 4, 5, 6, 7, 8, 9, 10, 11, 12 and 13 to be placed here

5 Conclusion

In this paper we have described a texture segmentation method which can be considered as a first step towards a completely data-driven segmentation algorithm. The proposed approach does not require any training on the textures to be partitioned. The number of regions in the image need not to be known either. The method makes use of local and global spatial statistics within a global statistical estimation framework relying on Markov Random Field models.

The method only requires the tuning of one parameter which determines the thresholds on the different texture features. Several examples have shown the capability of the model to yield usable segmentations on hand-assembled images consisting of natural textures as well as on complex real-world textured images.

Several significant extensions could be proposed for the segmentation procedure described in this paper. A completely data-driven segmentation algorithm could be developed by considering estimation techniques for the global parameter p . Statistical information criteria could be of some help here, although they are cpu time consuming [40, 42, 45]. Another point concerns the choice of the different texture features used in the segmentation. The same features have been considered for all images with good experimental results. Further improvements can be expected from an optimal selection of texture features for a given class of images. To

figure 6, a fine tuning of this parameter is not necessary to obtain a relevant segmentation (values between $0.45 \leq p \leq 0.55$ are adequate in this case). The degradation of the resulting segmentation is progressive when the value of p is tuned within the range $[0, 1]$ (figure 6d, 6e and 6f).

Similar qualitative results have been obtained on several other hand-assembled images corresponding to different mosaic patterns (figure 7, 8 and 9). For these different examples, the algorithm was able to recover the right number of regions, although the region boundaries are located only approximately in some cases. Besides the smoothing effect of the MRF prior, the lack of accuracy in the location of region boundaries is also related to the window sizes used in the local texture analysis. A minor adjustment of parameter p was required to process these different examples.

The processing of real-world (textured) images is generally more involved, mainly because the textures appearing in natural images are usually non stationary. We present results of segmentation on four real-world images, which have been selected for their representativity and their difficulty. The images show several textured regions with various (possibly non stationary) spatial statistics (figures 10, 11, 12, and 13). Artefacts or non-stationary areas in textures are difficult to handle and lead generally to over-segmented images, if they are not taken into account in the model. The approach proposed here, although using stationary texture models, seems quite robust to these phenomena, as will be seen in the results.

Figure 10 corresponds to an aerial image showing two cultivated fields : the boundary between the two regions is delineated quite accurately, although the upper region is not homogeneous. Figure 11 presents another aerial image. This kind of image is typical of real-world applications, in which the textures are non-homogeneous and where the number of regions is difficult to determine, even visually. In this case four significant regions are obtained : in particular, the forest area has been segmented in two regions, with different spatial patterns. The two other regions are cultivated fields.

Figure 12 shows a satellite image of the south-eastern part of France, with similar difficulties. The area includes a part of the Alps and a part of the Mediterranean Sea. Even on this hard example (which shows highly non-stationary textures), a relevant segmentation (Fig. 12b) is obtained. The algorithm enables to separate quite accurately the different significant areas : mountain, land and sea. A last example, corresponding to an ultrasound image of a kidney with a high level of speckle noise is presented figure 12. As can be seen in figure 12b, the algorithm extracts approximately the ob-

No training was performed on the textures and the number of regions was assumed unknown. The features related to the joint occurrence matrices were computed on 16 gray level images in 5x5 local windows, along 4 directions with a distance $d=3$. The other parameters (gray levels, contrast and directional residuals) were computed on a 15x15 window. Parameter μ was set to 1.5 which corresponds to a good balance between the a priori energy U_2 and the data term U_1 . The only parameter that needed to be tuned was parameter p (defined by Eqn. 8) which determines the different thresholds on the texture features. The results turned out to be largely insensitive to the adjustment of parameter p (the “optimal” value could be modified by 10% without noticeable changes in the segmentation : an example of the results obtained for different values of parameter p is presented in the following). The cpu time, on a Sun-4 station reached about 25 mn, for a 128x128 image.

We first describe the results obtained on various texture mosaics, including different textures and object geometry. Figure 4a presents a first mosaic, which is composed of three different textures placed against a background texture. Figure 4b shows the boundary of the underlying mosaic, which corresponds to the ground truth. The boundaries of the final segmentation are superimposed on the original image in figure 4c. The error between the ground truth and the segmentation obtained by the unsupervised method is depicted figure 4d. Four regions (including the background) are obtained in this case for a value $p = 0.5$. The boundaries of the different regions are delineated quite accurately, even if the MRF prior tends to “smooth” the region boundary in high curvature areas (the corners of the different rectangles for example). More sophisticated models (including for instance line processes and geometrical constraints on edges, [21]) could be introduced to cope with this problem.

Figure 5 shows the different intermediate steps of the relaxation process. The maximum likelihood estimate, used in the initialization step (Section 3.3) is presented in figure 5a. As can be seen, this initial estimate, which does not take into account the a priori term U_2 , is noisy and partial. The result of the first relaxation step appears in figure 5b : the a priori term introduced here clearly enables to constrain the solution and to smooth out noisy points. Figure 5c depicts an intermediate state of the label field, after relabeling the outlier class ρ . The final segmentation is shown in figure 5d.

Figure 6 illustrates, on the same texture sample, the sensitivity of the result with respect to the choice of parameter p . Let us recall that parameter p is a probability defined by Eqn. (8) ($0 \leq p \leq 1$). As can be seen on

corresponds to one single region in the image. At the end of each relaxation stage (i.e. when the label field is stable), the connected components of the outlier class are relabeled, creating new regions (figure 3). However, label ρ is maintained for all small connected components (of size less than 1 % of the image size). These points will not be considered for the computation of global region characteristics, but may be visited and relabeled in the next relaxation step. This procedure enables to let small region grow until they reach a sufficient size for providing significant statistics. The global statistical distributions associated to each significant region are updated after the relabeling process.

This relaxation-relabeling process is pursued until stability is obtained on the label field. In practice, small oscillations may be observed, and the process is stopped after three complete relaxation-relabeling procedures, with satisfactory results. The value of parameter ϕ , which controls the creation of new regions, is increased during these three iterations, in order to disfavour the creation of small artifacts (regions) at the end of the segmentation process (successive values of $\phi = 2.5, 3.5$ and 4.5 were adopted in all experiments).

Figure 3 to be placed here

Initialization

Since the relaxation algorithm is deterministic, the final segmentation map will depend on the initial label configuration. As was already stated, the initial value $\hat{R} = 1$ corresponds to one single region. This initial single region is updated using a maximum likelihood estimator of the label field, which is nothing but minimizing the global energy without the prior term U_2 . In this step, the value of parameter Φ (that controls the outlier class) has been set to a large value ($\Phi_0 = 10$, in all experiments). This large value disfavors noisy detections of small regions at the beginning of the segmentation process (this step defines the initial label configuration for the relaxation - relabeling steps previously described).

4 Experimental results

The segmentation technique described in the previous section has been applied to different mosaics composed of natural textures extracted from the Brodatz album [7], as well as on real-world textured images.

In the ICM algorithm [3], the label which produces the largest decrease of the energy function is selected to update site s . This amounts to the minimization of a local energy function which is derived from the global energy U by computing the energy variation induced by updating the site s [3, 21]. In deriving this local energy function, we assume that the global statistics associated to each region remain unchanged during the updating process. This is a good approximation in practice, provided that the different existing regions are large enough. This condition is guaranteed by filtering the small regions as explained in the next section.

To cope with the unknown number of regions, an additional label ρ , corresponding to an "outlier class" is introduced in the MRF modeling. The outlier label ρ is assigned to site s if the Kolmogorov distance between the local distribution of texture features on block B_s (centered at s) and the global distributions on all existing regions is too large. More precisely, between the labels of sites neighboring s and label ρ , the algorithm selects the label $r = \hat{r}$ which minimizes the following local energy (derived, as explained previously, from the global energy function U) :

For $r \neq \rho$:

$$\begin{aligned} \Delta U_s(r) = & \sum_{t \text{ neighbors of } s} \mu (1 - 2\delta(r - e_t)) \\ & + \sum_{i=1}^m \left\{ 2 \Gamma(d(o^{(i)}(R(e_s = r)), o^{(i)}(B_s)) > c^{(i)}) - 1 \right\}. \end{aligned} \quad (9)$$

The local energy associated to the outlier class $r = \rho$ is defined by :

$$\Delta U_s(\rho) = \sum_{c \in C} \mu (1 - 2\delta(\rho - e_t)) + \phi. \quad (10)$$

ϕ is the parameter of the outlier class ($\phi > 0$ and a high value for ϕ disfavors the creation of new regions). A value of $\phi = 2.5$ was adopted in all experiments. This local update is applied in turn to the different sites in the image, using an "instability heap" to focus first on sites with incorrect labels. This technique is inspired from the *HCF* method proposed by Chou *et al.* in [11].

Determination of R

The current number of regions \hat{R} is updated once the relaxation on the label field e has converged. An initial value of $\hat{R} = 1$ is used, which

$$V_1(o(B_s), o(R(e_s))) = \sum_{i=1}^m \left\{ 2 \Gamma(d(o^{(i)}(R(e_s)), o^{(i)}(B_s)) > c^{(i)}) - 1 \right\} \quad (7)$$

where $d(o^{(i)}(R(e_s)), o^{(i)}(B_s))$ stands for the Kolmogorov-Smirnov distance between the distribution of feature $o^{(i)}$ on the whole region labeled e_s , and the distribution of the same feature on block B_s . $c^{(i)}$ is a threshold which can be explicitly determined using statistical tables associated to the Kolmogorov limit distribution. Indeed this is related to the probability p that the Kolmogorov-Smirnov distance exceeds $c^{(i)}$ whereas the two distributions are the same :

$$p = \text{Prob} \left\{ d(o^{(i)}(R(e_s)), o^{(i)}(B_s)) > c^{(i)} \right\} \quad (8)$$

Function $\Gamma = 1$ if $d(o^{(i)}(R(e_s)), o^{(i)}(B_s)) > c^{(i)}$, $\Gamma = 0$ otherwise. This potential favors the labeling of site s with the region number $e_s = r$ if the local distribution at this site is consistent with the global distribution of the texture features in region r .

The minimization of the global energy function (Eqn.3), as well as the determination of the number of regions R is performed using a deterministic relaxation algorithm.

3.3 The relaxation algorithm

To minimize the global energy function, stochastic relaxation[21] is not considered here, since it is too involved from a computational point of view. We have adopted a modified version of the standard deterministic ICM algorithm proposed by Besag [3], which alternately updates the segmentation map e and the current number of regions.

Estimation of e

Let us denote by \hat{R} the current number of regions. The candidate region labels for a given site s are thus labels $r = 1, \dots, \hat{R}$. To save cpu time, we have only considered as candidate labels for site s the labels currently assigned to the neighbors of s (including e_s), without noticeable degradation on the results.

This energy function is defined as the sum of two contributions. The term U_2 describes the prior on the region label map e . This term is related to a Markov model associated to an eight-neighborhood system. The model favors spatially homogeneous regions, by the choice of potential functions of the form :

$$V_2(e_s, e_t) = \mu(1 - 2\delta(e_s - e_t)), \quad (4)$$

where μ is a positive parameter and $\delta(x) = 1$ for $x = 0$, $\delta(x) = 0$ otherwise. In this energy function, only the doubleton (two-element) cliques are considered (figure 1).

Figure 1 to be placed here

The global energy term U_2 is derived, according to [21] :

$$U_2(e) \triangleq \sum_{(s,t) \text{ neighbours}} V_2(e_s, e_t). \quad (5)$$

Term U_1 describes the interaction between observations and labels. It is inspired by a model described by Geman *et al.* [20], with several significant differences. In [20] the model is based on a Kolmogorov-Smirnov distance expressing the similarity of the distribution of texture features in blocks centered at two neighboring sites s and t . We introduce here a different potential function which measures the distance, at a given site s , between the distribution of texture features in a block B_s centered at s (local distribution) and the distribution of texture features in the whole candidate region to which we want to assign site s (global distribution) (figure 2). This technique enables us, in particular, to create new regions by detecting outliers during the relaxation process, as will be explained in section 3.3.

Accordingly, energy function U_1 is decomposed as :

$$U_1(o, e, R) = \sum_{s \in S} V_1(o(B_s), o(R(e_s))), \quad (6)$$

where $o(B_s) = \{o_r, r \in B_s\}$ is the set of feature vectors on block B_s centered at s and $o(R(e_s)) = \{o_r, r \in R(e_s)\}$ is the set of feature vectors belonging to the region labeled e_s (figure 2).

Figure 2 to be placed here

The potential function V_1 is defined by :

3 Unsupervised texture segmentation

The algorithm described in this paper is unsupervised in the following sense :

- it does not require *a priori* knowledge on the textures of the image (in particular no learning step is necessary) ;
- the number of regions or texture classes needs not to be known either ;
- only one parameter needs to be tuned in the segmentation procedure.

3.1 Statistical texture features

The observations, on which the segmentation process is based, correspond to vectors of texture features describing various spatial statistics. These features are computed on local windows and are assigned to the pixel located at the center of the window. The features which have been used include seven standard features proposed by Haralick (mean, variance, energy, correlation, entropy, contrast, and homogeneity), derived from the joint occurrence matrices (see [24]). These features are computed on 16 gray-level quantized images in 5x5 windows along 4 different directions. Three other features, which have been shown by Geman *et al.* to lead good texture discrimination [20] have been considered : the gray level, the local contrast, and an “isotropic residual” [20]. Other parameters have been tested, but have not included in the final version of the segmentation algorithm. The set of texture measures retained here has led to good and robust results on all the processed images.

To each site is thus assigned a feature vector $o_s = (o^1(s), \dots, o^m(s))$ where $o^i(s)$ designates the *i*th texture feature computed at site *s*.

3.2 The global energy function

As in Section 2, e_s , $s \in S$, will designate the hidden variable assigned to site *s*, which describes the label of the region to which site *s* belongs. $o = \{o_s, s \in S\}$ represents the observation field corresponding to vectors of statistical texture features.

The issue is to estimate field $e = \{e_s, s \in S\}$ and to determine *R* (the number of regions in the image) from the observations *o*. A MAP estimator is used to this end. The MAP estimate of *e* is obtained by minimizing the following global energy function :

$$\hat{e} = \arg \min_{e \in \Omega} U_1(e, o, R) + U_2(e). \quad (3)$$

$$\hat{e} = \arg \min_{e \in \Omega} U(o, e). \quad (1)$$

Due to the Markovian property of the field, the energy decomposes as follows [21] :

$$U(o, e) = \sum_{c \in \mathcal{C}} V_c(o, e), \quad (2)$$

where \mathcal{C} denotes the set of cliques associated to neighborhood system \mathcal{G} . Cliques c are subsets of sites which are mutual neighbors. The potential function V_c is locally defined on clique c and expresses the local interactions between the different variables of the clique (see [21] for an extensive theory of MRF). The form of the potential functions V_c is problem dependent and defines the local and global properties of the model.

Defining global energy functions is a powerful tool for specifying non-linear interactions between different image features (luminance, edges, region labels, etc.). They help to combine and organize spatial and temporal information by introducing strong generic knowledge about the features to be estimated. Energy minimization techniques have been used in a wide range of applications including image restoration [3, 21], edge detection [20], luminance and texture segmentation [15, 17, 20, 23, 36], stereovision [1], computed tomography, surface reconstruction [11], visual motion analysis [5, 6, 27, 34] and scene interpretation [37]. Energy functions involve generally two components, one of which expresses the interaction between the hidden labels and the observations and the other which encodes constraints on the desired solution [20]. The choice of these energy functions is either heuristic or may be guided by a statistical modeling of the interaction between the variables.

Minimizing the global energy function U is however usually a hard optimization problem : the number of possible label configurations is generally *very large* and, moreover, the global energy function U may contain local minima. Computationally demanding stochastic relaxation algorithms are therefore generally necessary to compute exact MAP solutions. Less cpu-time consuming deterministic relaxation algorithms such as ICM, [3], can often be used instead, when a good initial guess is available (other approaches include Graduated Non Convexity, Mean field Annealing or multiscale relaxation, [29]). In the following, we use a modified version of the ICM algorithm to determine jointly the region labels and the number of texture classes in the image. A good initial segmentation is obtained from a maximum likelihood estimator.

tion of texture features in the whole candidate region to which we want to assign the site s (global distribution). Comparing local and global distributions enables robust segmentation, even on complex real-world images, since global spatial statistics (obtained from regions) are used instead of features computed on local windows only as often proposed.

Besides, the described method does not require any training and the thresholds used for the discrimination between different texture classes are determined using a statistical table. The method only requires the tuning of one parameter which is a probability that determines all the thresholds on the texture features.

The remainder of this paper is organized as following : in Section 2 we briefly describe the general framework of the statistical approach used here (global bayesian estimation and MRF models). The segmentation algorithm is described in Section 3. Results on mosaics of natural textures, as well as real-world textured images are presented in Section 4. They show the ability of the model to discriminate textures and to recover the different regions in the image.

2 Bayesian image analysis and Markov Random Fields

The general mathematical framework under concern in this paper is based on global bayesian estimation along with MRF models. The goal of this theory is to extract a 2D field of discrete labels describing the information of interest in a given image processing task from a given set of observed variables. In texture segmentation, for instance, the labels to be estimated are region labels.

Let $O = (O_s, s \in S)$ designate an observation field defined on a rectangular lattice S . Let $E = (E_s, s \in S)$ denote the unobserved (hidden) label field, defined on the same lattice S ¹. Realizations of fields O (resp. E) will be denoted $o = \{o_s, s \in S\}$ (resp. $e = \{e_s, s \in S\}$). Let Λ be the (discrete) state space of random variable e_s and Ω the (finite) set of all possible label configurations $e = \{e_s, s \in S\}$. If (E, O) is assumed to be a Markov Random Field with respect to a chosen neighborhood system $\mathcal{G} = \{\mathcal{G}_s, s \in S\}$ it can be shown [21], that the best estimate \hat{e} of e given o (according to the Maximum A Posteriori (MAP) criterion) is obtained by minimizing a global energy function $U(o, e)$:

¹Different lattices for E and O can also be adopted, see [21]

image, with different region numbers are to be computed. This can be time consuming, especially when there are many regions.

Alternately, several authors [10, 20, 22] have proposed, within the same statistical framework (bayesian estimation and MAP criterion), to describe texture using standard statistical features which are computed in local windows. The use of local texture features, instead of global texture models (such as Gauss-Markov Models), reduces significantly the number of parameters to be determined. Many different types of features have been used in texture segmentation. These include joint occurency matrices [24], power spectrum measures, run lenght statistics [19], directional residuals [20] or characteristics obtained from mathematical morphology transforms [24]. In [20], various texture features are used, along with an MRF model of the segmentation map. The algorithm relies on a deterministic constrained relaxation on the region labels. It is partially unsupervised but the number of regions is assumed to be known approximately and thresholds for discriminating the different textures have to be fine-tuned. A two-class non supervised segmentation algorithm relying on joint occurency matrices and an MRF model is developed in [16] to guide a vision-equipped robot mower.

In this paper, we propose an alternative approach to unsupervised texture segmentation (a preliminary version of this method has been presented in [28]). The method relies on a global bayesian estimation framework. Following [20, 33], we describe textures using standard statistical features which are computed in local windows [24]. Besides, an MRF model is used to introduce global constraints on the computed segmentation map. The optimal segmentation is defined as the Maximum A Posteriori (MAP) estimate of the region labels given the texture features. The optimal segmentation is approximated using a deterministic relaxation procedure which simultaneously determines the current number of regions. This is achieved by introducing an additional state (corresponding to an outlier class) in the MRF model representing the image partition. This additional label enables the dynamic creation of new regions during the relaxation process (a similar idea has been suggested for motion-based segmentation by Bouthemy and Francois [5]). Mixed windows may thus be split in several sub-regions, leading to reliable and robust parameter estimation. The energy function associated to the MRF model is based on the Kolmogorov-Smirnov distance between local and global statistical distributions of texture features. The model measures the distance, at a given site s , between the distribution of texture features in a local window B_s centered at s (local distribution) and the distribu-

segmentation system based on the computation of localized histograms and region merging procedures is presented in [4]. In [33], an edge detection technique is developed to extract the boundaries between different textures in a multidimensional space of texture features. In [30] texture energy features are clustered using a K-means algorithm. The segmentation is refined by probabilistic relaxation techniques.

A more powerful and now widespread approach to “unsupervised” texture segmentation consists in modeling the different textures with stationary statistical models. Usual models include Gauss Markov Random Fields (GMRF) [12, 18, 36, 38, 43, 41] and binomial models [14, 31]. Non-linear Markov Random Field (MRF) models are often jointly used to constrain the image partition [22, 31, 35, 36, 38]. The image is first divided into disjoint local windows [10, 12, 18, 31, 36, 38, 41]. Each region is assumed to contain mainly one texture. This assumption enables to identify locally the parameters of stationary texture models [10, 12, 18, 38, 41]. Windows with nearby parameters are then merged using bayesian clustering techniques [10, 36, 41] or region growing procedures [18]. This first crude segmentation may afterwards be refined using relaxation techniques relying on global bayesian estimation and the Maximum A Posteriori (MAP) criterion [31, 36, 38] or Maximum Likelihood approaches [12]. In [43], a quadtree structure is alternately used to decompose the original image. Split and merge procedures, based on Gauss-Markov hypotheses on the underlying textures, are defined on the hierarchical structure.

One problem of these methods is the difficulty of the computation, from local windows, of reliable parameters for the underlying model. A second drawback is related to the processing of mixed windows (i.e. windows containing several different textures), which is generally quite involved. In these approaches, the cluster validation, i.e., the estimation of the number of textures present in the image is also a problem. [10].

In [35], Lakshmanan *et al.* describe an unsupervised segmentation scheme for luminance images, based on an MRF model. Model parameters are estimated alternately with the region labels, yielding a partial optimal solution. A similar technique is used by Won *et al.*, [42], for the segmentation of textures described by Gauss-Markov fields. The number of regions in the image is determined using a model fitting criterion (this technique was first proposed in [45] to estimate the number of classes in a grey-level image). However, when a model fitting criterion is used, several segmentations of the original

1 Introduction

Unsupervised image segmentation remains a fundamental issue in low-level image analysis and computer vision since most high-level processing schemes must rely on reliable and robust segmentations. The purpose of segmentation is to partition the image into meaningful components. Each component corresponds to a region of similar attribute (luminance, color, texture, depth, motion, etc.) [5, 15, 23]. The extracted components may be used in a second step in medium or high-level identification or interpretation schemes.

In this paper we address the problem of unsupervised texture segmentation. This problem is known to be more involved than the segmentation of luminance, primarily because texture is not directly observed (the adaptation of the method described in this paper to luminance or color segmentation is straightforward). The proposed segmentation scheme relies on local texture features in combination with a global statistical model describing the statistics of the segmentation map.

Various approaches, models and techniques have been proposed in the literature for the analysis, classification, or segmentation of textured images [8, 9, 23, 24, 25]. Since the early works of Julesz [32] and Haralick [26, 24], significant improvements have been obtained in texture segmentation by introducing global statistical models [15, 17, 20, 23, 36]. Markov Random Fields (MRF) [17], in particular, define a class of statistical models which enable to describe both the local and global properties of textures and segmentation maps. Segmentations of high quality have been obtained using these models, provided that the parameters of the model and the number of regions are known. On the other hand, unsupervised texture segmentation remains an intricate issue, although the problem has received considerable attention in the last decade [16, 20, 31, 33, 36, 38, 42]. The segmentation of textured images is indeed usually “semi-supervised” in the sense that, there is a supervised selection of texture characteristics, or an *a priori* knowledge of the number of texture classes has to be introduced [23]. Completely data-driven segmentation methods, able to handle real-world textured images, remain beyond the present state-of-the-art.

Early attempts to non supervised texture segmentation have made use of texture measures (features derived from cooccurrence matrices for instance) associated to merging or edge detection techniques. A hypothesis testing technique for merging texture features is described by Connors *et al.* [13] to segment high resolution aerial images. A general purpose unsupervised

Segmentation non supervisée des textures par une approche markovienne

Résumé : La segmentation des images texturées constitue une étape préliminaire cruciale dans de nombreuses applications en analyse d'images. Les approches par modélisation statistique conduisent à de bons résultats dans ce domaine, lorsque les paramètres des modèles statistiques et le nombre de régions à extraire sont connus *a priori*. La segmentation non supervisée d'images texturées reste, par contre, un problème délicat, auquel aucune solution complète n'a été apportée jusqu'à présent. Nous contribuons à cet effort, en proposant une méthode de segmentation ne nécessitant pas de connaissance *a priori* sur le nombre ou le type de textures présentes dans l'image. L'algorithme de segmentation s'appuie sur l'analyse des statistiques spatiales des images originales. Un modèle markovien est utilisé pour spécifier les propriétés d'homogénéité de la carte de segmentation recherchée. Une estimation bayésienne de cette carte est obtenue grâce à un algorithme de relaxation déterministe, qui détermine conjointement le nombre de régions dans l'image. Un seul paramètre gère l'ensemble de la procédure de segmentation. Des résultats sur des mosaïques de textures naturelles et sur des exemples d'images réelles texturées démontrent l'aptitude du modèle à produire des segmentations de bonne qualité, sans connaissance *a priori* sur les caractéristiques des composantes texturées de l'image.

Mots-clé : analyse d'images, segmentation non supervisée, analyse de textures, modèles statistiques, champs markoviens



A Markov Random Field model-based approach to unsupervised texture segmentation using local and global spatial statistics

Charles Kervrann, Fabrice Heitz

Programme 4 — Robotique, image et vision
Projet TEMIS

Rapport de recherche n° 2062 — Octobre 1993 — 31 pages

Abstract: The general problem of unsupervised textured image segmentation remains a fundamental but not entirely solved issue in image analysis. Many studies have proven that statistical model-based texture segmentation algorithms yield good results provided that the model parameters and the number of regions be known *a priori*. In this paper, we present an unsupervised texture segmentation method which does not require a priori knowledge about the different texture regions, their parameters, or the number of available texture classes. The proposed algorithm relies on the analysis of local and global second and higher order spatial statistics of the original images. The segmentation map is modeled using an augmented-state Markov Random Field, including an outlier class which enables dynamic creation of new regions during the optimization process. A bayesian estimate of this map is computed using a deterministic relaxation algorithm. The whole segmentation procedure is controlled by one single parameter. Results on mosaics of natural textures and real-world textured images show the ability of the model to yield relevant and robust segmentations when the number of regions and the different texture classes are not known *a priori*.

Key-words: image analysis, non-supervised texture segmentation, statistical models, Markov Random Field models

(Résumé : *tsvp*)

**A Markov Random Field model-based approach
to unsupervised texture segmentation using local
and global spatial statistics**

Charles Kervrann, Fabrice Heitz

N° 2062

Octobre 1993

PROGRAMME 4

Robotique,
image
et vision

 *Rapport
de recherche*

1993